

## **Prueba de comparación de mentira dirigida híbrida (HDLC) de EyeDetect**

### **Resumen de desarrollo y validación**

16 de mayo de 2021

Recientemente Converus lanzó la prueba de comparación de mentira dirigida híbrida (HDLC) de EyeDetect. El protocolo HDLC es una prueba para incidentes específicas que cubre un solo tema relevante, como un asesinato, robo, o consumo de drogas. El tema de comparación para este tema relevante (R1) es una pregunta de mentira dirigida. Las preguntas de mentira dirigida se refieren a actos inapropiados, y la computadora indica al sujeto de la prueba que mienta a esas preguntas. Algunos ejemplos son: "En toda su vida, ¿alguna vez ha mentido para salir de problemas?" o "En toda su vida, ¿siempre ha dicho la verdad?" A los sujetos se les pide que mientan a esas preguntas, que afirmen que nunca han mentido y que siempre han dicho la verdad. La computadora informa al sujeto que utilizará los datos fisiológicos y óculo motores para tratar de distinguir sus respuestas engañosas a preguntas de mentira dirigida de las supuestas respuestas veraces a preguntas relevantes. A los sujetos se les dice que si la computadora no puede distinguir entre la mentira dirigida y las preguntas relevantes, se le dice al sujeto que podría reprobar la prueba. El protocolo HDLC asume que los sujetos veraces estarán preocupados de no aprobar la prueba porque no reaccionan como deberían a las preguntas de mentira dirigida. En consecuencia, la prueba predice que los sujetos veraces estarán más preocupados por las preguntas de mentira dirigida que por las preguntas relevantes. Por otro lado, la prueba predice que los sujetos culpables reaccionarán con más fuerza a las preguntas relevantes porque están mintiendo a esas preguntas que pertenecen al asunto que se está investigando. Para los sujetos que responden a preguntas relevantes con sinceridad, las preguntas con mentiras dirigidas deberían representar la mayor amenaza para aprobar la prueba y captar más atención. Por el contrario, para los sujetos de prueba que mientan a preguntas relevantes, las preguntas relevantes son más destacadas y deberían representar la mayor amenaza.

El protocolo HDLC combina medidas tradicionales de polígrafo y medidas óculo motoras más nuevas para distinguir entre personas veraces y engañosas. Además de las señales óculo motoras del rastreador de ojos, EyeDetect+ 2.0 graba la actividad electro dérmica de electrodos desechables conectados al superficie palmar de los dedos de la mano no dominante. También graba la respiración de un medidor de tensión alrededor del pecho o el abdomen, el electrocardiograma (ECG) de los electrodos conectados a cada brazo y el foto pletismógrafo (PPG) de un dedo de la mano no dominante.

A partir de las señales de ECG y PPG, la computadora obtiene una medida indirecta de la presión arterial conocida como tiempo de tránsito del pulso (PTT). El PTT varía inversamente con los cambios en la presión arterial (Geddes et al., 1981; Obrist et al., 1978; Obrist et al., 1979). A medida que aumenta la presión arterial, disminuye el tiempo que tarda el pulso en viajar desde el corazón a la periferia. La investigación ha demostrado que el PTT es al menos tan diagnóstico como el cardiógrafo tradicional para pruebas de polígrafo de mentira probable (Webb y Kircher, 2005). EyeDetect+ 2.0 reemplaza el electrocardiograma con PTT. El PTT es menos invasivo que el electrocardiograma, alivia la incomodidad del sujeto y ofrece más flexibilidad en la redacción de pruebas.

Desarrollamos y validamos de forma cruzada dos modelos de decisiones. El primer modelo asumió la disponibilidad de medidas de polígrafo y óculo motoras. Los resultados no fueron concluyentes cuando el puntaje de credibilidad estaba entre 45 y 55. En la validación cruzada, las decisiones de HDLC fueron 88,6% correctas, excluyendo el 8,1% de los resultados inconclusos. La precisión fue ligeramente mayor para los inocentes (89,3%) que para los culpables (87,9%). Sin una región inconcluso, la precisión media fue del 86,2%. La precisión fue ligeramente mayor para los veraces (87,1%) que para los culpables (85,4%).

El segundo modelo supuso la disponibilidad de solo medidas óculo -motoras. En la validación cruzada, el 89,5% de los sujetos inocentes y el 84,5% de los culpables se clasificaron correctamente. La precisión media fue del 87,0%, excluyendo el 7,3% de los resultados inconclusos. Sin una región inconclusa, las decisiones de HDLC fueron 82,3% y 87,1% correctas para sujetos inocentes y culpables, respectivamente. La precisión media fue del 84,7%.

El resto del presente artículo describe los procedimientos utilizados para desarrollar la prueba HDLC y estimar su precisión.

### **Experimento**

Realizamos un experimento de crimen simulado siguiendo el modelo de Cook, et al. (2012) para recopilar los datos fisiológicos necesarios para desarrollar y validar de forma cruzada un modelo de decisión compuesto por medidas de polígrafo y óculo motoras. Reclutamos a 124 sujetos de prueba de la comunidad local. A los sujetos se les dijo que algunos participantes cometerían un robo simulado de \$20 de la billetera de una secretaria, mientras que otros serían inocentes y no cometerían el crimen. Los sujetos fueron asignados aleatoriamente a grupos culpables e inocentes. Los sujetos culpables (n=60) tomaron \$20 en efectivo de la oficina de una secretaria y mintieron al respecto en la prueba. Los sujetos inocentes (n=60) respondieron con sinceridad a las preguntas sobre el robo de los \$20.

Los sujetos llevaron a cabo su tarea asignada y se les aplicó la prueba HDLC. El supervisor de la pruebas le pidió al sujeto que se sentara en un escritorio frente a una pantalla de computadora. El supervisor de pruebas adjuntó sensores de conductancia cutánea, ECG, foto pletismógrafo (PPG) y respiración al sujeto y luego calibró el rastreador ocular. Usamos un monitor de fisiología multicanal para registrar la conductancia de la piel y la respiración a 340 Hz y ECG y PPG a 1000 Hz. Un rastreador de ojos remoto Tobii 4C conectado a la parte inferior del monitor de la computadora registró la posición de la mirada bilateral y el diámetro de la pupila a 60 Hz. Antes de que pudieran tomar la prueba HDLC, los sujetos debían demostrar que entendían la instrucción de mentir a las preguntas de mentira dirigida. Una voz digital de texto a voz presentó instrucciones de prueba y preguntas por auriculares. Se indicó a los sujetos que evitaran moverse, que mantuvieran los pies apoyados en el suelo y que miraran la pantalla de la computadora durante la prueba.

La prueba HDLC consistía en un protocolo optimizado para las medidas tradicionales de polígrafo seguido de un protocolo óculo motor optimizado para medidas óculo motoras. La fase optimizada del polígrafo de la prueba consistió en un conjunto de ocho preguntas de tipo Sí / No que se repitieron tres veces con intervalos de 22 segundos entre preguntas. A esto le siguieron tres repeticiones de un

conjunto de 36 enunciados de tipo Verdadero / Falso optimizados para las medidas óculo motoras. Se les pidió a los sujetos que respondieran lo enunciados de tipo Verdadero / Falso con la mayor rapidez y precisión posible o podrían fallar en la prueba. El análisis realizado a la mitad de la recopilación de datos indicó que los datos de la fase optimizada para las medidas óculo motoras no aumentaron la precisión lograda en la fase inicial optimizada del polígrafo. En ese momento, abandonamos la fase de optimización ocular y aumentamos el número de sesiones de polígrafo de tres a cuatro. Al final de la recopilación de datos, comparamos la precisión diagnóstica lograda con tres o cuatro sesiones de polígrafo. La precisión lograda con cuatro sesiones de polígrafo no fue más alta que la obtenida con solo tres sesiones de polígrafo. En consecuencia, la prueba HDLC final contenía un conjunto de ocho preguntas de tipo Sí / No que se repitieron tres veces.

Una voz digital de texto a voz presentó instrucciones de prueba y preguntas por auriculares. Se indicó a los sujetos que evitaran moverse, que mantuvieran los pies apoyados en el suelo y que miraran la pantalla de la computadora durante la prueba.

La Tabla 1 contiene las preguntas presentadas en una de las tres sesiones de HDLC. La computadora presentó la misma pregunta de mentira dirigida y preguntas relevantes en las tres sesiones, pero el orden de presentación de las preguntas y las preguntas aritméticas variaron durante las sesiones.

*Tabla 1. Preguntas de la prueba de comparación híbrida de mentira dirigida*

No. de pregunta	Tipo de pregunta	Pregunta
1	Mentira dirigida	En toda su vida, ¿alguna vez ha roto alguna regla?
2	Relevante	¿Se llevó esos \$20 en efectivo?
3	Aritmética	¿6 + 2 es igual a 8?
4	Relevante	¿Tienes esos \$20 en efectivo contigo ahora?
5	Mentira dirigida	En toda su vida, ¿alguna vez cometió un error?
6	Aritmética	¿5 + 1 es igual a 6?
7	Mentira dirigida	En toda su vida, ¿siempre ha dicho la verdad?
8	Relevante	¿Sacó los \$20 en efectivo de la mochila?

A los sujetos se les pagó \$40 en efectivo por participar en el estudio y se les ofreció un bono adicional de \$30 si aprobaron la prueba.

### **Análisis**

Se analizaron datos óculo motores para identificar características que discriminaran entre preguntas y declaraciones respondidas de manera veraz y engañosa. Creamos una ecuación de regresión logística a partir de un subconjunto de características del polígrafo diagnóstico y del motor ocular. La ecuación de regresión logística proporcionó la probabilidad de que el sujeto de la prueba fuera veraz en las

preguntas relevantes de la prueba (puntuación de credibilidad). La puntuación de credibilidad se utilizó para clasificar al sujeto de la prueba como veraz o engañoso. En algunos casos, el puntaje de credibilidad estuvo cerca del 50% (probabilidad) y se clasificó como inconcluso. Los resultados de precisión se reportan con y sin una región inconclusa.

### **Validación de N-veces**

Un modelo estadístico que es óptimo para clasificar los casos en un experimento en particular rara vez es óptimo para la población de la que se tomaron muestras de los sujetos. El modelo no es óptimo porque la muestra no representa perfectamente la población general de la que se extrajo. En consecuencia, el modelo produce estimaciones de precisión sesgadas si se prueba en los casos utilizados para crear el modelo.

Se pueden obtener mejores estimaciones de precisión con la validación de "N" veces. Una validación de N veces divide el conjunto de datos en N subconjuntos. El primer subconjunto comprende una submuestra de reserva y se elimina del conjunto de datos. Los subconjuntos restantes se combinan para crear un conjunto de entrenamiento. Se desarrolla un modelo de regresión logística utilizando los casos del conjunto de entrenamiento. Luego, ese modelo de regresión se utiliza para clasificar los casos en la submuestra reservada. La precisión observada en la muestra reservada proporciona una estimación de precisión menos sesgada porque los casos reservados no se utilizaron para optimizar los coeficientes de características en la ecuación de regresión. Se registra la precisión alcanzada en la muestra retenida.

Este proceso continúa para cada partición del conjunto de datos. El primer subconjunto se devuelve al conjunto de entrenamiento y el segundo subconjunto se elimina para que sirva como una nueva muestra reservada. Se crea una nueva ecuación de regresión logística con todos menos el segundo subconjunto de casos. Ese modelo se utiliza para clasificar casos en la muestra reservada y se registra su precisión. Este proceso se repite para cada uno de los subconjuntos restantes. La mejor estimación de la precisión del modelo es la precisión media para las k muestras reservadas.

### **Validación del HDLC con polígrafo y medidas óculo motoras**

Se realizó una validación de 4 veces en el modelo de decisión. La muestra de 124 sujetos se dividió en 4 submuestras de 30 sujetos cada una. La cuarta submuestra contenía 34 sujetos. La mitad de los sujetos de cada submuestra eran inocentes y la otra mitad culpables. La Tabla 2 muestra el porcentaje correcto de preguntas veraces y engañosas para cada pliegue, así como la precisión media en los cuatro pliegues.

Tabla 2. Porcentaje de decisiones correctas para sujetos inocentes y culpables en una validación cuádruple

	1ra vez	2da vez	3ra vez	4ta vez	Media
<b>n</b>	30	30	30	34	
<b>Inocente</b>	94.1	86.7	80.0	86.7	87.1
<b>Culpable</b>	76.5	73.3	93.3	100.0	85.5
			<b>Precisión media</b>		<b>86.3</b>

La precisión fue ligeramente mayor para los inocentes (87,1%) que para los culpables (85,5%). Con base en estos resultados, esperaríamos que el HDLC produzca un 86% de decisiones correctas cuando el modelo se utilice en una nueva muestra.

Debido a que la prueba HDLC es más probable que se use en investigaciones criminales que la RCT o MCT, puede ser difícil justificar la toma de una decisión definitiva cuando el puntaje de credibilidad está cerca de 50 (probabilidad). Usamos las puntuaciones de credibilidad no sesgadas para cada muestra reservada de 30 casos para calcular las tasas de precisión cuando las pruebas que produjeron puntuaciones de credibilidad entre 45 y 55 se consideraron inconclusas. Los resultados se presentan en la Tabla 3. De las 124 pruebas de HDLC, 10 (8,1%) fueron inconclusos. El porcentaje de decisiones correctas en la Tabla 3 excluye los resultados inconclusos.

Tabla 3. Frecuencia de resultados correctos, incorrectos y inconclusos para sujetos inocentes y culpables

	Tamaño de la muestra	Correcto	Incorrecto	Inconcluso	% Decisiones correctas
<b>Inocente</b>	62	50	6	6	89.3
<b>Culpable</b>	62	51	7	4	87.9
	124	101	13	<b>Precisión Media</b>	<b>88.6</b>

Cuando se prueba en una nueva muestra de casos, esperaríamos que la prueba HDLC produzca aproximadamente un 89% de decisiones correctas, excluyendo un 8% de resultados de prueba inconclusos.

### Validación del HDLC con solo medidas óculo motoras

Para obtener medidas de polígrafo tradicionales, se deben conectar varios sensores al sujeto y asegurarse de que los distintos sensores produzcan señales de alta calidad. Este contacto directo con el sujeto y entrenamiento adicional del supervisor de pruebas que no es necesario con una prueba de engaño óculo motora que utiliza un rastreador de ojos remoto.

Se desarrolló un nuevo modelo de HDLC y se realizó una validación cruzada que solo utilizaba medidas oculares-motoras. Los resultados de la validación cuádruple se presentan en la Tabla 4.

*Tabla 4. Porcentaje de decisiones correctas para sujetos inocentes y culpables en una validación de 4 veces utilizando solo medidas óculo motoras*

	1ra vez	2da vez	3ra vez	4ta vez	Media
<b>n</b>	30	30	30	34	
<b>Inocente</b>	86.7	66.7	86.7	88.2	82.3
<b>Culpable</b>	80.0	100.0	86.7	82.4	87.1
			<b>Precisión media</b>		<b>84.7</b>

La precisión media de HDLC con solo medidas oculares-motoras fue del 84,7% en la validación cruzada. Obtuvimos los resultados en la Tabla 5 con una región no concluyente para puntajes de credibilidad insesgados que oscilaron entre 45 y 55. De las 124 pruebas, 9 no fueron concluyentes (7.3%). Excluyendo los resultados inconclusos, la precisión media fue del 87%. La adición de medidas de las señales de polígrafo tradicionales aumentó la precisión del 87% a casi el 89%.

*Tabla 5. Frecuencia de resultados correctos, incorrectos y no concluyentes para sujetos inocentes y culpables usando solo medidas oculares-motoras*

	Tamaño de la muestra	Correcto	Incorrecto	Inconcluso	% Decisiones correctas
<b>Inocente</b>	62	51	6	5	89.5
<b>Culpable</b>	62	49	9	4	84.5
	124	100	15	<b>Precisión media</b>	<b>87.0</b>

## Referencias

- Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S., Woltz, D. J., & Kircher, J. C. (2012). Lin' Eyes: Ocular-motor Measures of Reading Reveal Deception. *Journal of Experimental Psychology: Applied*, *18*(3), 301-313.
- Geddes L.A., Voelz, M.H., Babbs C.F., Bourland J.D., & Tacker W.A. (1981). Pulse transit time as an indicator of arterial blood pressure. *Psychophysiology*, *18*, 71–74.
- Obirst, P. A., Light, K. C., McCubbin, J. A., Hutcheson, J. S., & Hoffer, J. L. (1978). Pulse transit time: Relationship to blood pressure. *Behavior Research Methods & Instrumentation*, *10*, 623-626.
- Obirst, P. A., Light, K. C., McCubbin, J. A., Hutcheson, J. S., & Hoffer, J. L. (1979). Pulse transit time: Relationship to blood pressure and myocardial performance. *Psychophysiology*, *16*, 292-301.
- Webb, A. K. & Kircher, J. C. (2005). *Use of Pulse Transit Time for the Psychophysiological Detection of Deception*. Final report to the U.S. Department of Defense. Salt Lake City: University of Utah, Department of Educational Psychology.